Mutant DNA encoding insulin receptor substrate 1

Abstract
A DNA sequence encoding insulin receptor substrate 1 (IRS-1), the DNA sequence containing a mutation of at least one nucleotide, and the protein encoded by said DNA sequence.

Inventors: Pedersen; Oluf (Holte, DK); Bj.o slashed.rb.ae butted.k; Christian (Frederiksberg C, DK); Frederiksen; Kathrine Almind (K.o slashed.benhavn .O slashed., DK)
Assignee: Novo Nordisk A/S (Bagsvaerd, DK)
Appl. No.: 557139
Filed: February 12, 1996
PCT Filed: June 10, 1994
PCT NO: PCT/DK94/00227
371 Date: February 12, 1996
102(e) Date: February 12, 1996
PCT PUB.NO.: WO94/29345
PCT PUB. Date: December 22, 1994

Foreign Application Priority Data

Jun 10, 1993[DK] 0683/93
Aug 09, 1993[DK] 0915/93

Current U.S. Class: 435/320.1; 435/325; 530/350; 536/23.1; 536/23.5
Intern'l Class: C12N 015/00; C12N 005/00; C07K 001/00; C07H 021/04
Field of Search: 800/2,DIG. 1-4 435/325,320.1 536/23.1,23.5

References Cited [Referenced By]

U.S. Patent Documents
5162215 Nov., 1992 Bosselman et al. 435/172.
5545808 Aug., 1996 Hew et al. 800/2.
5621075 Apr., 1997 Kahn et al. 530/350.
Foreign Patent Documents
92/13083 Aug., 1992 WO.

Other References
Shoelson et al. YMXM motifs of IRS-1 define substrate specificity of the insulin receptor kinase. PNAS, vol. 89, No. 6, pp. 2027-2031, Mar. 1992.
Pursel et al. Genetic engineering of livestock. Science, vol. 244, pp. 1281-1288, Jun. 16, 1989.
Houdebine, L. M. Production of pharmaceutical proteins from transgenic animals. J. of Biotechnology, vol. 34, pp. 269-287, 1994.
Dialog Information Services, file 155, Medline, Dialog accession No. 0852900, Medine accession No. 93239000.
Dialog Information Services, file 155, Medline, Dialog accession No. 08058050, Medline accession No. 92196050.
Dialog Information Services, file 155, Medline, Dialog accession No. 08231886, Medline accession No. 92369886.
Dialog Information Services, file 155, Medline, Dialog accession No. 0726027, Medline accession No. 90167273.
Dialog Information Services, file 155, Medline, Dialog accession No. 07601382, Medline accession No. 91120382.
Almind, K., et al., J. Clin. Invest., vol., 97, No. 11, pp. 1-7 (1996).
Dialog Information Services, file 155, Medline, Dialog accession No. 08680176, Medline accession No. 93390176.

Primary Examiner: Chambers; Jasemine C.
Assistant Examiner: Clark; Deborah J. R.
Attorney, Agent or Firm: Zelson, Esq.; Steve T., Gregg, Esq.; Valeta

Parent Case Text

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of PCT/DK94/00227 filed Jun. 10, 1994, which is incorporated herein by reference.

Claims

We claim:

1. An isolated DNA sequence comprising the sequence of SEQ ID NO:1 encoding insulin receptor substrate 1 (IRS-1), the DNA sequence containing a mutation of at least one nucleotide selected from the group consisting of a mutation of G to A in the first position of codon 972, a mutation of A to G in the third position of codon 805, and a mutation of G to C in the third position of codon 894, wherein the mutation interferes with signal transduction through IRS-1.

2. The DNA sequence of claim 1, wherein the mutation gives rise to at least one amino acid substitution in the IRS-1 sequence.

3. The DNA sequence of claim 1 comprising a mutation of G to A at nucleotide 3494 of SEQ ID NO:1.

4. A recombinant expression vector comprising the DNA sequence of claim 1.

5. An isolated mammalian cell containing and expressing the DNA sequence of claim 3.

6. An isolated IRS-1 protein containing at least one amino acid substitution wherein glycine is substituted by arginine at position 972 in SEQ ID NO:2 and wherein said substitution interferes with signal transduction through IRS-1.

Description

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of PCT/DK94/00227 filed Jun. 10, 1994, which is incorporated herein by reference.

FIELD OF INVENTION

The present invention relates to a mutant DNA sequence encoding insulin receptor substrate 1, a method of detecting a mutation in the gene encoding insulin receptor substrate 1, as well as a diagnostic composition and a test kit for use in the method.

BACKGROUND OF THE INVENTION

Non-insulin-dependent diabetes mellitus (NIDDM) is a common endocrine disorder and a considerable body of evidence strongly suggests that genetic factors contribute to the pathogenesis (1). Studies of patients with overt NIDDM and of individuals at high risk of NIDDM reveal abnormalities of both insulin secretion and insulin action (2). However, a genetic defect at one or more loci in the cellular action of insulin and insulin-like growth factor 1 (IGF1) might well involve both the biochemical pathways of tissues that regulate insulin secretion and insulin sensitive hepatic and extrahepatic tissues that produce or extract glucose.

Although more than twenty different mutations of the insulin receptor gene have been reported in syndromes of severe insulin resistance frequently associated with the skin disorder acanthosis nigricans or ovarian hyperandrogenism (17), mutations in the insulin receptor molecule do not explain the genetic etiology of the common form of NIDDM.

The common form of late onset NIDDM is a heterogeneous disorder where at least two major defects contribute to the pathophysiology of the phenotype: insulin resistance and insulin deficiency (2). Most likely NIDDM is also polygenic and it is suggested that subsets of patients will display changes in various

genes, in aggregate accounting for the inherited components of the disorder. The high cumulative risk of diabetes in offspring of NIDDM parents (30-50%) and the high concordance rate in identical twins (70-100%) underscore the significance of the genetic etiology of the disease (1).

Insulin initiates its cellular effects by binding to the a subunit of its tetrameric plasma membrane receptor (3). The kinase in the .beta. subunit is thereby activated which in turn catalyzes the intramolecular autophosphorylation of specific tyrosine residues of the .beta. subunit, further stimulating the tyrosine kinase activity of the receptor towards other protein substrates in the cascade of insulin action.

Recently, the first endogeneous substrate for the insulin receptor kinase (termed insulin receptor substrate 1, abbreviated to IRS-1) was cloned and sequenced (4-6). The complementary DNA sequence encodes a cytoplasmic, hydrophilic protein of a relative molecular mass between 165 and 185 kD (27,28) which contains multiple phosphorylation sites. Besides being a substrate for the insulin receptor kinase, IRS-1 is also phosphorylated following the activation of the IGF1 receptor kinase (16).

IRS-1 is barely detectable in cells expressing few insulin receptors, but is strongly detected in cells expressing high levels of receptors and weakly detected in cells expressing mutant receptors defective in biological signalling (27,28). IRS-1 is a unique molecule containing 20 tyrosine phosphorylation consensus sequences, 6 of which appear in YMXM (Tyr-Met-X-Met) motifs. Following insulin stimulated tyrosine phosphorylation of YMXM motifs in the IRS-1 molecule, the phosphorylated IRS-1 binds phosphatidylinositol 3-kinase (PI3-kinase) suggesting that IRS-1 acts as a multisite "docking" protein to bind signal proteins thereby linking the receptor kinase to insulin sensitive transporters and enzymes (7-15). The PI3-kinase is composed of at least two subunits including a 110 kDa catalytic subunit and a 85 kDa regulatory protein which contains src homology 2 domains that mediate protein-protein interactions by binding to phosphotyrosine residues in various proteins (7-15). Interestingly, it has been demonstrated that insulin causes the interaction between IRS-1 and PI-3 kinase via phosphorylated YMXM motifs of IRS-1 and src homology 2 domains of the 85 kDa regulatory subunit of PI-3 kinase. Moreover, overexpression of IRS-1 potentiates the activation of PI3-kinase in insulin stimulated cells, and tyrosyl phosphorylated IRS-1 or synthetic peptides containing phosphorylated YMXM motifs activate PI3-kinase in vitro. Besides being a substrate of the insulin receptor kinase, IRS-1 is also phosphorylated following the activation of the IGF1 receptor kinase (16). Hence, it has been suggested that IRS-1 by binding and regulating enzymes containing src homology 2 domains may play a critical role to select and differentiate the effects of insulin and IGF1 from those of other tyrosine kinases and to generate diversity and amplification of signal transmission into multiple intracellular pathways (7-16).

SUMMARY OF THE INVENTION

According to the present invention, it has surprisingly been found that a number of NIDDM patients carry mutations in the gene coding for IRS-1. It is at present assumed that one or more of the mutations may be involved in or associated with the etiology of NIDDM, and their presence may therefore be diagnostic for NIDDM and possibly also other disorders resulting from insulin resistance.

Accordingly, the present invention relates to a DNA construct comprising a DNA sequence encoding insulin receptor substrate 1 (IRS-1), the DNA sequence containing a mutation of at least one nucleotide, or comprising a fragment of the DNA sequence including said mutation.

It is at present assumed that mutation of the IRS-1 gene may be indicative of abnormalities significant for the development of NIDDM or other disorders. For instance, the mutation may give rise to the substitution of an amino acid in IRS-1 which may cause changes in the tertiary structure of IRS-1. Such changes may interfere with the normal interaction between the insulin or IGF-1 receptor kinase and one or more intracellular proteins regulating cellular metabolism and growth. These proteins typically have src homology 2 domains and include phosphatidyl inositol 3 kinase, GRB-2 and SHPTP-2 (vide M. F. White et al., Exp. Clin. Endocrinol. 101, Suppl. 2, 1993, pp. 98-99, in particular FIG. 1). Mutations may also interfere with the transcription or translation of the gene, or with the stability of the IRS-1 transcript. Alternatively, the mutation may be associated with (i.e. genetically linked with) the mutation which causes the disease.

In another aspect, the present invention relates to a living system containing a DNA construct of the invention and capable of expressing IRS-1 wherein at least one amino acid is substituted. The living system, which may comprise a cell or a multicellular organism containing the appropriate signal

transduction pathway, may be used to screen for substances which have an effect on insulin or IGF-1 stimulated signal transduction in the system.

In a further aspect, the present invention relates to a method of detecting the presence of a mutation in the gene encoding IRS-1, the method comprising obtaining a biological sample from a subject and analysing the sample for a mutation of at least one nucleotide. Based on current knowledge, it is assumed that the present method may be used to diagnose predisposition to NIDDM in a subject as well as other disorders resulting from insulin resistance (such as android obesity, essential hypertension, dyslipidemia or atherosclerosis). Biological samples may, for instance, be obtained from blood, serum, plasma or tissue. The invention further relates to a diagnostic composition and a test kit for use in the method.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows mutation screening using single stranded conformation polymorphism technique (SSCP) of nucleotides 3338-3600 of the human insulin receptor substrate 1 (IRS-1 gene). The fragment was PCR-amplified from genomic DNA using specific oligonucleotide primers and .sup.32 P-labeled and subjected to nondenaturing gel electrophoresis as described in Methods. The autoradiogram shows the migration profiles of single stranded (ss DNA) as well as double stranded DNA (ds DNA). Lane 3 depicts the migration profile of an individual who is heterozygous (He) for the glycine.sup.972 mutation. Lanes 1 and 2 show the migration profile for the corresponding wild type (Wt).

FIG. 2 shows direct nucleotide sequencing of a part of the 3338-3600 base pair fragment of IRS-1. The sequencing was performed on the noncoding strand (c.f. Table 3). Upper panel shows the wild type sequence whereas lower panel depicts the nucleotide sequence from an individual who is heterozygous for a single base substitution in the coding strand at nucleotide position 3494 as indicated by an arrow, (G.fwdarw.A) (c.f. Table 3). The base substitution in codon.sup.972 caused a substitution of glycine with arginine.

DETAILED DESCRIPTION OF THE INVENTION

In particular, the present invention relates to a DNA construct comprising a DNA sequence encoding IRS-1 and containing a mutation giving rise to at least one amino acid substitution in the IRS-1 protein sequence. The mutation may for instance be located at a site where the amino acid substitution interferes with signal transduction through IRS-1, as such a mutation is most likely to be involved in disease etiology. An example of such a DNA sequence is one containing a mutation of G to A in the first position of codon 972 of the IRS-1 gene.

This mutation leads to substitution of glycine by arginine in position 972. The molecular mechanism by which the IRS-1 gene variant may lead to development of NIDDM is not known at present. However, compared with glycine, arginine is a much larger molecule and has a polar side chain with a positive charge and a high pKa of 12.5. In IRS-1, codon 972 is located between two YMXM motifs. Although the changes which the glycine for arginine substitution may cause in the tertiary structure of IRS-1 are not known at present, it is assumed that the steric and electrostatic changes of the IRS-1 mutant interfere with the normal interaction between insulin and IGF-1 mediated phosphorylation of neighbouring YMXM motifs of IRS-1 and signal transmitting proteins with src homology 2 domains. Alternatively, the mutation may be a marker associated with another mutation in this or another gene, which other mutation is the one actually involved in disease etiology. This may also be the case with the silent mutations found in the third position of codon 805 (A to G) and in the third position of codon 894 (G to C).

Clinical investigations have shown that NIDDM patients as a whole have insulin resistant glucose disposal to peripheral tissues when compared with glucose-tolerant control subjects. However, glycine.sup.972 -mutation carriers with NIDDM did not differ in their degree of insulin resistance when compared with mutation-negative diabetic patients. Furthermore, the sensitivity of peripheral tissues to insulin was measured in 2 of 3 mutation-positive nondiabetics who turned out to have values of insulin sensitivity within the normal range. In contrast, all mutation carriers had remarkably low fasting plasma concentrations of insulin and C-peptide when compared with matched noncarriers. The combination of comparable insulin sensitivity of peripheral tissues and low basal levels of circulating .beta.-cell secretory products may indicate pancreatic .beta.-cell dysfunction.

In a preferred embodiment, the DNA construct of the invention comprises the DNA sequence shown in the Sequence Listing as SEQ ID NO:1, or a fragment of said DNA sequence including the mutation of G to A at

nucleotide 3494 of SEQ ID NO:1.

The length of the DNA construct may vary widely depending on the intended use. For use as an oligonucleotide probe for hybridisation purposes, the DNA fragment may be as short as 17 nucleotides. For expression in a living system as defined above, the DNA construct will typically comprise the full-length DNA sequence encoding IRS-1.

The DNA construct of the invention comprising the mutation in the DNA sequence encoding IRS-1 may suitably be of genomic or cDNA origin, for instance obtained by preparing a genomic or cDNA library and screening for DNA sequences coding for all or part of the IRS-1 by hybridization using synthetic oligonucleotide probes in accordance with standard techniques (cf. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989). The probes used should be specific for the mutation. Alternatively, the DNA sequence encoding wild-type IRS-1 may be modified by site-directed mutagenesis using synthetic oligonucleotides containing the mutation for homologous recombination in accordance with well-known procedures. The DNA sequence may also be prepared by polymerase chain reaction using specific primers, for instance as described in U.S. Pat. No. 4,683,202, or Saiki et al., Science 239, 1988, pp. 487-491.

The DNA construct of the invention may also be prepared synthetically by established standard methods, e.g. the phosphoamidite method described by Beaucage and Caruthers, Tetrahedron Letters 22 (1981), 1859-1869, or the method described by Matthes et al., EMBO Journal 3 (1984), 801-805. According to the phosphoamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed and ligated. This procedure may preferably be used to prepare fragments of the IRS-1 encoding DNA sequence.

The recombinant expression vector into which the DNA construct is inserted may be any vector which may conveniently be subjected to recombinant DNA procedures. The choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated (e.g. a viral vector).

In the vector, the mutant DNA sequence encoding IRS-1 should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the mutant DNA encoding IRS-1 in mammalian cells are the SV40 promoter (Subramani et al., Mol. Cell Biol. 1 (1981), 854-864), the MT-1 (metallothionein gene) promoter (Palmiter et al., Science 222 (1983), 809-814) or the adenovirus 2 major late promoter.

The mutant DNA sequence encoding IRS-1 may also be operably connected to a suitable terminator, such as the human growth hormone terminator (Palmiter et al., op. cit.). The vector may further comprise elements such as polyadenylation signals (e.g. from SV40 or the adenovirus 5 Elb region), transcriptional enhancer sequences (e.g. the SV40 enhancer) and translational enhancer sequences (e.g. the ones encoding adenovirus VA RNAs).

The recombinant expression vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. An example of such a sequence is the SV40 origin of replication. The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the gene coding for dihydrofolate reductase (DHFR) or one which confers resistance to a drug, e.g. neomycin, hygromycin or methotrexate.

The procedures used to ligate the DNA sequences coding for IRS-1, the promoter and the terminator, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., op.cit.).

In a further aspect, the present invention relates to a variant of IRS-1 containing at least one amino acid substitution, in particular a variant containing at least one amino acid substitution at a site where the substitution interferes with signal transduction through IRS-1, or a fragment thereof including said substitution. An example of such a variant is one in which glycine.sup.972 is substituted by arginine, or a

fragment thereof containing said substitution, e.g. the variant which has the amino acid sequence shown in the Sequence Listing as SEQ ID NO:2, or a fragment thereof containing Arg.sup.972.

The living system into which the DNA construct of the invention is introduced may be a cell which is capable of producing IRS-1 and which has the appropriate signal transduction pathways. The cell is preferably a eukaryotic cell, such as a vertebrate cell, e.g. a Xenopus laevis oocyte or mammalian cell, in particular a mammalian cell. Examples of suitable mammalian cell lines are the COS (ATCC CRL 1650), BHK (ATCC CRL 1632, ATCC CCL 10), CHL (ATCC CCL39) or CHO (ATCC CCL 61) cell lines. Methods of transfecting mammalian cells and expressing DNA sequences introduced in the cells are described in e.g. Kaufman and Sharp, J. Mol. Biol. 159 (1982), 601-621; Southern and Berg, J. Mol. Appl. Genet. 1 (1982), 327-341; Loyter et al., Proc. Natl. Acad. Sci. USA 79 (1982), 422-426; Wigler et al., Cell 14 (1978), 725; Corsaro and Pearson, Somatic Cell Genetics 7 (1981), 603, Graham and van der Eb, Virology 52 (1973), 456; and Neumann et al., EMBO J. 1 (1982), 841-845.

The mutant DNA sequence encoding IRS-1 may then be expressed by culturing a cell as described above in a suitable nutrient medium under conditions which are conducive to the expression of the IRS-1-coding DNA sequence. The medium used to culture the cells may be any conventional medium suitable for growing mammalian cells, such as a serum-containing or serum-free medium containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection).

The living system according to the invention may also comprise a transgenic animal. A transgenic animal is one in whose genome a heterologous DNA sequence has been introduced. In particular, the transgenic animal is a transgenic non-human mammal, mammals being generally provided with appropriate signal transduction pathways. The mammal may conveniently be a rodent such as a rat or mouse. The mutant DNA sequence encoding IRS-1 may be introduced into the transgenic animal by any one of the methods previously described for this purpose. Briefly, the DNA sequence to be introduced may be injected into a fertilised ovum or cell of an embryo which is subsequently implanted into a female mammal by standard methods, resulting in a transgenic mammal whose germ cells and/or somatic cells contain the mutant DNA sequence. For a more detailed description of a method of producing transgenic mammals, vide B. Hogan et al., Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. The mutant DNA sequence may also be introduced into the animal by transfection of fertilised ova with a retrovirus containing the DNA sequence, cf. R. Jaenisch, Proc. Natl. Acad. Sci. USA 73, 1976, pp. 1260-1264. A further method of preparing transgenic animals is described in Gordon and Ruddle, Methods Enzymol. 101, 1983, pp. 411-432.

In one embodiment of the present method of detecting the presence of a mutation in the IRS-1 gene, a biological sample is obtained from a subject, DNA (in particular genomic DNA) is isolated from the sample and digested with a restriction endonuclease which cleaves DNA at the site of the mutation, and cleavage of the DNA within the gene encoding IRS-1 at this site is determined. After digestion, the resulting DNA fragments may be subjected to electrophoresis on an agarose gel. DNA from the gel may then be blotted onto a nitrocellulose filter and hybridised with a radiolabelled probe. The probe may conveniently contain a DNA fragment of the IRS-1 gene spanning the mutation (substantially according to the method of E. M. Southern, J. Mol. Biol. 98, 1975, pp. 503, e.g. as described by B. J. Conner et al., Proc. Natl. Acad. Sci. USA 80, 1983, pp. 278-282).

In a variant of this embodiment, the DNA isolated from the sample may be amplified prior to digestion with the restriction endonuclease. Amplification may suitably be performed by polymerase chain reaction (PCR) using oligonucleotide primers based on the appropriate sequence of IRS-1 spanning the site(s) of mutation, essentially as described by Saiki et al., Science 230, 1985, pp. 1350-1354. After amplification, the amplified DNA may be digested with the appropriate restriction endonuclease and subjected to agarose gel electrophoresis. The restriction pattern obtained may be analysed, e.g. by staining with ethidium bromide and visualising bands in the gel by means of UV light. As a control, wild-type DNA encoding IRS-1 (i.e. not containing the mutation) may be subjected to the same procedure, and the restriction patterns may be compared.

In the method of the invention, the sample is preferably analysed for a mutation located at a site where amino acid substitution interferes with signal transduction through IRS-1. A specific example of such a mutation is the mutation of G to A in the first position of codon 972 of the gene encoding IRS-1. This mutation results in a new restriction endonuclease cleavage site

5'-C C A/T G G-3' (SEQ ID NO:3)

3'-G G T/A C C-5' (SEQ ID NO:3) in the mutant DNA sequence coding for IRS-1.

An example of a suitable restriction endonuclease is BstN1.

A further embodiment of the method of the invention is an adaptation of the method described by U. Landegren et al., Science 241, 1988, pp. 1077-1080, which involves the ligation of adjacent oligonucleotides on a complementary target DNA molecule. Ligation will occur at the junction of the two oligonucleotides if the nucleotides are correctly base paired.

In a still further embodiment of the present method, the DNA isolated from the sample may be amplified using oligonucleotide primers corresponding to segments of the gene coding for IRS-1. The amplified DNA may then be analysed by hybridisation with a labelled oligonucleotide probe comprising a DNA sequence corresponding to at least part of the gene encoding IRS-1 and containing a mutation of at least one nucleotide, which mutation corresponds to the mutation the presence of which in the gene encoding IRS-1 is to be detected. As a control, the amplified DNA may furthermore be hybridised with a labelled oligonucleotide probe comprising a DNA sequence corresponding to at least part of the wild-type gene encoding IRS-1. This procedure is, for instance, described by DiLella et al., Lancet 1, 1988, pp. 497-499. Another PCR-based method which may be used in the present invention is the allele-specific PCR method described by R. Saiki et al., Nature 324, 1986, pp. 163-166, or D. Y. Wu et al., Proc. Natl. Acad. Sci. USA 86, 1989, pp. 2757-2760, which uses primers specific for the mutation in the IRS-1 gene.

Other methods of detecting mutations in DNA are reviewed in U.

Landegren, GATA 9, 1992, pp. 3-8. A currently preferred method of detecting mutations is by single stranded conformation polymorphism (SSCP) analysis substantially as described by Orita et al., Proc. Natl. Acad. Sci. USA 86, 1989, pp. 2766-2770, or Orita et al., Genomics 5, 1989, pp. 874-879.

The label substance with which the probe is labelled is preferably selected from the group consisting of enzymes, coloured or fluorescent substances, or radioactive isotopes.

Examples of enzymes useful as label substances are peroxidases (such as horseradish peroxidase), phosphatases (such as acid or alkaline phosphatase), .beta.-galactosidase, urease, glucose oxidase, carbonic anhydrase, acetylcholinesterase, glucoamylase, lysozyme, malate dehydrogenase, glucose-6-phosphate dehydrogenase, .beta.-glucosidase, proteases, pyruvate decarboxylase, esterases, luciferase, etc.

Enzymes are not in themselves detectable but must be combined with a substrate to catalyse a reaction the end product of which is detectable. Examples of substrates which may be employed in the method according to the invention include hydrogen peroxide/tetramethylbenzidine or chloronaphthole or o-phenylenediamine or 3-(p-hydroxyphenyl) propionic acid or luminol, indoxyl phosphate, p-nitrophenylphosphate, nitrophenyl galactose, 4-methyl umbelliferyl-D-galactopyranoside, or luciferin.

Alternatively, the label substance may comprise coloured or fluorescent substances, including gold particles, coloured or fluorescent latex particles, dye particles, fluorescein, phycoerythrin or phycocyanin.

In a particularly favoured embodiment, the probe is labelled with a radioactive isotope. Radioactive isotopes which may be used for the present purpose may be selected from I-125, I-131, In-111, H-3, P-32, C-14 or S-35. The radioactivity emitted by these isotopes may be measured in a beta- or gamma-counter or by a scintillation camera in a manner known per se.

For use in the present method, the invention further relates to a test kit for detecting the presence of a mutation in the gene encoding IRS-1, the kit comprising

(a) a restriction endonuclease which cleaves DNA at the site of the mutation,

(b) a first DNA sequence corresponding to at least part of the wild-type gene encoding IRS-1, and/or

(c) a second DNA sequence corresponding to at least part of the gene encoding IRS-1 and containing a mutation of at least one nucleotide, which mutation corresponds to the mutation the presence of which in

the gene encoding IRS-1 is to be detected.

The first DNA sequence may, for instance, be obtained from genomic DNA or cDNA encoding IRS-1 obtained from a healthy subject. The second DNA sequence may conveniently be a DNA construct according to the invention.

For use in the present method, the invention further relates to a test kit for detecting the presence of a mutation in the gene encoding IRS-1, the kit comprising

(a) means for amplifying DNA, and

(b) a labelled oligonucleotide probe comprising a DNA sequence corresponding to at least part of the gene encoding IRS-1 and containing a mutation of at least one nucleotide, which mutation corresponds to the mutation the presence of which in the gene encoding IRS-1 is to be detected.

Appropriate means for amplifying DNA (typically genomic DNA isolated from the biological sample) include, for instance, oligonucleotide primers, appropriate buffers and a thermostable DNA polymerase.

The invention is further illustrated in the following example which is not intended in any way to limit the scope of the invention as claimed.

EXAMPLES

Example 1

Subjects

A total of 86 NIDDM patients and 76 control subjects were included in the protocol. All study participants were unrelated Danish Caucasians and their clinical characteristics are given in Table 1 and 2. The control subjects had normal fasting plasma glucose levels and no family history of known diabetes. Patients with NIDDM, as defined by the National Diabetes Data Group (18), were recruited consecutively and unselected from the outpatient clinic at Steno Diabetes Center. All NIDDM patients were treated with diet, oral hypoglycemic drugs, or both. None of the participants in the study suffered from liver or kidney diseases as evaluated by clinical and standard laboratory examinations, and no subject was taking any other medication known to influence pancreatic .beta.-cell function or energy metabolism. Before participation, the purpose and risks of the study were carefully explained to all volunteers and their informed consent was obtained. The protocol was approved by the local Ethical Committee of Copenhagen and was in accordance with the Helsinki declaration.

Euglycemic, Hyperinsulinemic Clamp

NIDDM patients and 19 glucose-tolerant healthy control subjects were examined using a euglycemic, hyperinsulinemic clamp. The experiments were undertaken in the fasting state between 08.00 and 15.00 after a 10 h overnight fast. Each clamp comprised a 2 h basal period followed by a 4 h hyperinsulinemic glucose clamp. Details of the clamp technique have been described previously (19). To assess total peripheral glucose uptake, (3-.sup.3 H) glucose was infused throughout the study period. In the control subjects, (3-.sup.3 H) glucose was administered as a primed (25-.mu.Ci) continuous (0.25-.mu.Ci/min) infusion, whereas in the NIDDM subjects, the priming was increased in proportion to the increase in fasting plasma glucose concentration; the continuous infusion of labeled glucose was the same as in the control subjects (0.25 .mu.Ci/min). The clamp was performed by continuous infusion of 2 mU insulin.times.kg.sup.-1 .times.min.sup.-1 (Actrapid, Novo Nordisk, Denmark), and euglycemia was maintained by a variable infusion of 20% glucose at a rate determined by the measurement of plasma glucose levels at 5 to 10 min invervals. Total glucose disposal rate was calculated from the plasma concentrations of (3-.sup.3 H) glucose and plasma glucose with Steele's non-steady state equations (20). In these calculations, the distribution volume of glucose was taken as 200 ml/kg body weight and the pool fraction as 0.65. At the highest steady state plasma insulin level, where the hepatic glucose production is presumed to be nil, glucose infusion rates were used to calculate glucose disposal. Total peripheral glucose uptake was corrected for urinary glucose loss.

Preparation and Amplification of Genomic DNA

Genomic DNA was isolated from human leucocyte nuclei isolated from whole blood by protein kinase K digestion followed by phenol extraction on a Applied Biosystems 341 Nucleic Acid Purification System. A primary 50 .mu.l PCR reaction was carried out with 0,3 .mu.g of genomic DNA. The assay conditions were: 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM MgCl.sub.2, 0.1% Triton X-100, 0.2 mM dNTP's, 0.2 .mu.M of each oligonucleotide primer and 1.25 u Taq DNA polymerase (Promega, Madison, Wis.). Specific oligonucleotide primers for the human IRS-1 gene (4), 20-25 mers and 2 G's or C's at the 3' end, were synthezised on an Applied Biosystems 394 DNA/RNA syntheziser (Applied Biosystems Inc., Foster City, Calif.) and eluted on a NAB-10 column (Pharmacia P-L Biochemicals Inc, Milwaukee, Wis.) and used without further purification.

The nucleotide sequences of the DNA primers used for the polymerase chain reactions were as follows:

1.	553	5'GCTCAGCGTTGGTGGTGGCGGTGG3'	577 (SEQ ID NO: 4)
2.	783	3'CGACGAAGTTGTAGTTGTTCGCCCG5'	807 (SEQ ID NO: 5)
3.	727	5'GAAGTGGCGGCACAAGTCGAGCGC3'	756 (SEQ ID NO: 6)
4.	999	3'CGAGGCCGGAACCACTCCGACC5'	1020 (SEQ ID NO: 7)
5.	924	5'ACCGTGCTAAGGGCCACCACGACG3'	947 (SEQ ID NO: 8)
6.	1180	3'CCTCCGTCGCCGGCACCACGA5'	1200 (SEQ ID NO: 9)
7.	1128	5'TCTACCGCCTTTGCCTGACCAGC3'	1150 (SEQ ID NO: 10)
8.	1392	3'CGGTCAGGAGCAGGTTGACG5'	1411 (SEQ ID NO: 11)
9.	1337	5'ACCATCCTGGAGGCCATGCGG3'	1357 (SEQ ID NO: 12)
10.	1598	3'GGTCGGAGCCACCTGCCGTCGGGA5'	1621 (SEQ ID NO: 13)
11.	1545	5'CAGGCTCCTTCCGTGTCCGCG3'	1565 (SEQ ID NO: 14)
12.	1807	3'GGGTGCCGCTAGATCACGAAG5'	1827 (SEQ ID NO: 15)
13.	1757	5'CTGTCGTCCAGTAGCACCAGTGG3'	1779 (SEQ ID NO: 16)
14.	2007	3'GGGACTGGCGGGGGTTGCCAG5'	2037 (SEQ ID NO: 17)
15.	1953	5'GCGGTGAGGAGGAGCTAAGC3'	1932 (SEQ ID NO: 18)
16.	2200	3'GGGCAGGGTCAGGAGTCACCG5'	2230 (SEQ ID

NO: 19)

17. 2151 5'AGAGAACTCACTCGGCAGGC3'
2170 (SEQ ID NO: 20)

18. 2401 3'GGTGTGCCTACTACCGATGTACGG5'
2424 (SEQ ID NO: 21)

19. 2352 5'ACCCCTTGGAGCGTCGGGGG3'
2371 (SEQ ID NO: 22)

20. 2600 3'GGACTGAATCCTCCACCGGGGTCG5'
2623 (SEQ ID NO: 23)

21. 2546 5'CAGAGAGTGGACCCCAATGG3'
2566 (SEQ ID NO: 24)

22. 2800 3'CCTGAGGTTGTGGTCGTCGG5'
2819 (SEQ ID NO: 25)

23. 2712 5'TCTTGCCTCACCCCAAACCC3'
3150 (SEQ ID NO: 26)

24. 2964 3'CGAGACCAGCGGAAGAGATA5'
2983 (SEQ ID NO: 27)

25. 2918 5'GAGCCGGAGGAGGGTGCCCG3'
2938 (SEQ ID NO: 28)

26. 3180 3'GGTTCCGGTCGTGGAATGGA5'
3199 (SEQ ID NO: 29)

27. 3131 5'CAGACCAATAGCCGCCTGGC3'
3150 (SEQ ID NO: 30)

28. 3392 3'CCGTGACTCCTCATGTACTT5'
3411 (SEQ ID NO: 31)

29. 3339 5'CTTCTGTCAGGTGTCCATCC3'
3358 (SEQ ID NO: 32)

30. 3582 3'CGATGCACCTGTGGAGCGGT5'
3601 (SEQ ID NO: 33)

31. 3532 5'GGGCAGTGCCCAGCAGCCGG3'
3541 (SEQ ID NO: 34)

32. 3743 3'CGACGGGTGAGCAGGGACGACC5'
3764 (SEQ ID NO: 35)

33. 3687 5'CCTCAGCAGCCTCTGCTTCC3'
3712 (SEQ ID NO: 36)

34. 3950 3'CGTCGTCATCCCCCGCCACC5'
3969 (SEQ ID NO: 37)

35. 3900 5'CCACACCCAGTGCCACCCGG3'
3919 (SEQ ID NO: 38)

36. 4141 3'CCTGAAGTTTGTCACGGGAG5'
4160 (SEQ ID NO: 39)

37. 4067 5'GAGCCAGCCAAACTGTGTGG3' 4086 (SEQ ID NO: 40)

38. 4320 3'CCTGTAGTGTCGTCCAGCAA5' 4339 (SEQ ID NO: 41)

The mixture was overlaid with 40 .mu.l of mineral oil and after initial denaturation at 95.degree. C. for 3 min, the samples were subjected to 35 cycles of amplification: annealing at 60.degree. C. for 1 min, extension at 72.degree. C. for 2 min and denaturation at 94.degree. C. for 1 min. In these primary PCR reactions, fragments of 800-900 bp, were amplified. A secondary PCR reaction was performed with 1 .mu.l of a 1:10 diluted primary PCR product as template. The amplification conditions were as described above, except that (.alpha.-.sup.32 P)dCTP (Amersham International, Buckinghamshire, UK) was added. Each fragment amplified during secondary PCR had a size of 250-285 bp.

Single Stranded Conformation Polymorphism (SSCP) Analysis

In the primary mutation screening, SSCP analysis (according to the method of Orita et al., Genomics 5, 1989, pp. 874-879) of the entire coding region of IRS-1 was performed on genomic DNA from 19 insulin resistant NIDDM patients and 5 control subjects (Table 1). In all cases, 2 .mu.l of a secondary PCR product was combined with 8 .mu.l of a sequencing stop solution. One .mu.l was loaded unheated to identify double-stranded products and after heating at 94.degree. C. for 5 min, 2 .mu.l was loaded to the same wells on a 38.times.31.times.0,03 cm 5% polyacrylamid gel (49:1, acrylamid:bisacrylamid) in 90 mM Tris-borate, 2,5 mM EDTA, with 1% or 5% glycerol, respectively. In each patient SSCP analysis was undertaken at 2 different experimental conditions: electrophoresis was carried out with 1% glycerol at 35 W constant power for 4-5 h at 4.degree. C. by placing gel, buffer and electrophoresis apparatus at 4.degree. C. overnight, and with 5% glycerol at 65 W constant power for 2-3 hours at 25.degree. C. (21). The gels were transferred to 3 MM filter paper, covered with a plastic wrap and autoradiographed at -80.degree. C. with intensifying screens for 2-10 h. Gels contained 2 lanes with a known plasmid mutation and the corresponding wild type fragment as positive controls. In our laboratory the ability to detect known plasmid mutations (single base substitutions, small insertions, or small deletions) is 80-90% (22).

Synthesis of Single-Stranded DNA and Direct Nucleotide Sequencing

Single stranded DNA for sequencing was generated using biotinylated oligonucleotide primers and streptavidin-coated magnetic beads (available from Dynal AS, Oslo, Norway). Products were precipitated with ammoniumacetate and isopropanol and resuspended in appropiate volumes of water. Sequencing using the AutoRead Sequencing Kit (Pharmacia, Uppsala, Sweden) and Fluore-dATP was analysed on an automated laser fluorescence DNA sequencer (Pharmacia, Uppsala, Sweden).

Direct Restriction Enzyme Digestion of PCR Products

Restriction enzyme digestion was carried out in 15 .mu.l reactions, containing 10 .mu.l precipitated secondary PCR product, 1,5 .mu.l 10.times. NE buffer 2, 100 .mu.g/ml bovine serum albumin and 10 u of the restriction enzyme BstN 1 (New England BioLabs Inc., Beverly, Mass.). The fragments were analyzed on a 4.5% high solution agarose gel and visualized after staining with ethidium bromide.

Synthesis of cDNA

Percutaneous muscle biopsies (.about.500 mg) were obtained under local anesthesia from the vastus lateralis muscle of the thigh. Muscle samples were blotted to remove blood and connective and adipose tissue and were within 30 sec frozen in liquid N.sub.2 and stored at -80.degree. C. until assayed. Total RNA was isolated from muscle biopsies as described (23). cDNA was synthesized in 25 .mu.l volumes containing 1 .mu.g of total RNA, 0.2 mM deoxynucleoside triphosphates, 40 u RNasin (Promega, Madison, Wis.), 0.625 .mu.g oligo (dT).sub.18, 400 u Moloney Murine Leukaemia Virus Reverse Transcriptase (Life Technologies Inc., Grand Island, N.Y.), 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl.sub.2 and 10 mM DTT. The reactions were performed at 37.degree. C. for 1 hour, followed by enzyme inactivation by incubation for 10 min at 95.degree. C. SSCP was performed on cDNA as described for genomic DNA.

Other Assays

Plasma for analysis of chemical quantities was drawn from all study participants in the morning after 10 h of overnight fast. Plasma glucose was measured in duplicate by a hexokinase method. Tritiated glucose in plasma was determined as described (24). HbA.sub.1C was measured using a HPLC method (normal range 4.1-6.4%). Plasma insulin and C-peptide levels were analyzed by radioimmuno-assays (25,26).

Statistical Analysis

Data in text and tables are given as means.+-.SE. For comparisons Mann-Whitney's test for unpaired data was applied.

Primary Screening for Mutations in the IRS-1 Gene

In 19 overlapping fragments, each consisting of 250-285 bp, the entire coding region of the IRS-1 gene was analyzed using genomic DNA from 19 NIDDM patients and 5 control subjects. Compared with matched control subjects the NIDDM patients were as a group hyperinsulinemic ($p<0.02$) and were characterized by impaired insulin stimulated glucose disposal to peripheral tissues ($p<0.0001$) (Table 1). SSCP scanning showed 3 different aberrant migration profiles. Subsequent nucleotide sequencing revealed one polymorphism at codon 805 (GCA.fwdarw.GCG (alanine)). Four of 19 NIDDM patients were heterozygous for this single base substitution which did not predict any change in the amino acid sequence. Another silent mutation was detected at codon 894 (CCG.fwdarw.CCC (proline)). Only one of 19 NIDDM patients was heterozygous for this nucleotide base variation. However, at codon 972 SSCP analysis showed that three of 19 NIDDM patients were heterozygous for a mutation in which glycine (GGG) was substituted for arginine (AGG) (FIGS. 1 and 2). The mutation was confirmed by sequencing of both DNA strands and by direct enzymatic digestion of PCR products with the restriction enzyme Bst N1, for which a restriction site was created by the nucleotide substitution (data not shown). None of the 5 control subjects, who primarily were SSCP scanned, showed evidence of polymorphisms. The glycine 972 mutation was located between 2 Tyr-Met-X-Met motifs in the IRS-1 gene (Table 3). Subsequently, the codon 972 mutation, primarily identified on genomic DNA isolated from blood cells, was verified by studies of cDNA, synthesized from total RNA which was isolated from skeletal muscle biopsies.

Secondary Screening for Glycine.sup.972 Mutation

Using SSCP molecular scanning and specific enzymatic digestion of primary PCR products further 67 NIDDM patients (Table 2) and 76 healthy control subjects were examined for the occurrence of the glycine.sup.972 mutation. Seven of the 67 NIDDM patients were heterozygous for the mutation. Moreover, 3 of 76 healthy volunteers were also heterozygous carriers of the codon.sup.972 mutation. All 3 control subjects who were positive for the mutation had a normal glucose response to an oral glucose challenge (data not shown).

Clinical Characterization of Individuals Carrying the Glycine.sup.972 Mutation

Table 1 shows the results obtained from the clinical investigations of the 3 glucose-tolerant controls who were positive for the glycine.sup.972 mutation. Insulin-glucose clamp was performed in 2 of the 3 mutation carriers and the glucose disposal rates of these individuals were within the range of the 19 mutation-negative control subjects. However, interestingly also glucose- tolerant mutation carriers were characterized by relatively low values of fasting plasma insulin (decreased by 33-46%) and C-peptide (decreased by 25-40%), respectively, when compared with mean values obtained in mutation-negative healthy controls.

Table 2 gives a summary of the phenotypical characteristics of the 10 NIDDM patients who are heterozygous carriers of the glycine.sup.972 mutation in the IRS-1 gene when compared to 76 NIDDM patients who are negative for the same mutation. No significant differences were shown in age, known diabetes duration, body mass index, fasting levels of plasma glucose, HbA.sub.1C and basal or insulin stimulated glucose disposal rates. However, fasting plasma levels of insulin and C-peptide were reduced by 44% ($p<0.05$) and 37% ($p<0.02$), respectively, in diabetics who were positive for the glycine.sup.972 mutation when compared with mutation-negative diabetic subjects.

Example 2

A random sample of 383 unrelated healthy young Caucasians (15-32 years of age) was studied to determine whether the glycine.sup.972 mutation confers insulin resistance. All individuals had their

insulin sensitivity estimated using Bergman's minimal model (intravenous injections of glucose combined with tolbutamide), and the occurrence of the glycine.sup.972 mutation was determined by means of SSCP scanning (as described in Example 1) and restriction enzyme analysis of genomic DNA (as described in Example 1). 35 subjects were found to carry the mutation.

Those carriers of the glycine.sup.972 mutation who had a body mass index (BMI) of more than 25 kg/m.sup.2 (n=10, BMI=28.4.+-.0.9 kg/m.sup.2) (mean.+-.standard error) had a two-fold lower insulin sensitivity and a significantly higher fasting serum level of C-peptide than the non-carriers with the same BMI (n=99, BMI=28.1.+-.0.3 kg/m.sup.2). In contrast, there was no difference in the measured variables between the carriers and non-carriers of the mutation in subjects with a BMI of less than 25 kg/m.sup.2. In multivariate analysis adjusting for differences in VO.sub.2 max, BMI, gender and age, the presence of the glycine.sup.972 mutation within the group of obese subjects was negatively associated with insulin sensitivity ($p<0.04$) and positively associated with fasting serum C-peptide ($p<0.09$), fasting serum triglyceride ($p<0.02$) and serum total cholesterol ($p<0.03$).

The results of the study show that in young obese subjects, the glycine.sup.972 mutation is associated with whole-body insulin resistance and dyslipidemia.

REFERENCES

1. Hitman GA, McCarthy MI. Genetics of non-insulin-dependent diabetes mellitus. Baillieres Clin Endocrinol Metab 1991; 5:455-476.

2. DeFronzo RA. The triumverate: .beta.-cell, muscle, liver: a collision responsible for NIDDM. Diabetes 1988; 37:667-687.

3. Kahn CR, White MF. The insulin receptor and the molecular mechanism of insulin action. J Clin Invest 1988; 82:1151-1156.

4. Sun XJ, Rothenberg P, Kahn CR, et al. Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. Nature (Lond) 1991; 352:73-77.

5. Rothenberg P, Lane WS, Karasik A, Backer J, White MF, Kahn CR. Purification and partial sequence analysis of pp185, the major cellular substrate of the insulin receptor tyrosine kinase. J Biol Chem 1991; 266:8302-8311.

6. Nishiyama M, Wands JR. Cloning and increased expression of an insulin receptor substrate-1-like gene in human hepatocellular carcinoma. Biochem Biophys Res 1992; 183:280-285.

7. Backer JM, Schroeder GG, Kahn CR, et al. Insulin stimulation of phosphatidylinositol 3-kinase activity maps to insulin receptor regions required for endogenous substrate phosphorylation. J Biol Chem 1992; 267:1367-1364.

8. Myers Jr MG, Backer JM, Sun XJ, et al. IRS-1 activates phosphatidylinositol 3-kinase by associating with src homology 2 domains of p85. Proc Natl Acad Sci USA 1992; 89:10350-10354.

9. Backer JM, Myers Jr MG, Shoelson SE, et al. Phosphatidylinositol 3-kinase is activated by association with IRS-1 during insulin stimulation. The EMBO Journal 1992; 11: 3469-3479.

10. Yonezawa K, Ueda H, Hara K, et al. Insulin-dependent formation of a complex containing an 85-kDa subunit of phosphatidylinositol 3-kinase and tyrosine-phosphorylated insulin receptor substrate 1. J Biol Chem 1992; 267:25958-25966.

11. Folli F, Saad MJA, Backer JM, Kahn CR. Insulin stimulation of phosphatidylinositol 3-kinase activity and association with insulin receptor substrate 1 in liver and muscle of the intact rat. J Biol Chem 1992; 267:22171-22177.

12. Yonezawa K, Yokono K, Shii K, et al. In vitro association of phosphatidylinositol 3-kinase activity with the activated insulin receptor tyrosine kinase. J Biol Chem 1992; 267:440-446.

13. Hayashi H, Kamohara S, Nishioka Y, et al. Insulin treatment stimulates the tyrosine phosphorylation

of the .alpha.-type 85-kDa subunit of phosphatidylinositol 3-kinase in vivo. J Biol Chem 1992; 267:22575-22580.

14. Sun XJ, Miralpeix M, Myers Jr MG, et al. Expression and function of IRS-1 in insulin signal transmission. J Biol Chem 1992; 267:22662-22672.

15. Shoelson SE, Chatterjee S, Chaudhuri M, White MF. YMXM motifs of IRS-1 define substrate specificity of the insulin receptor kinase. Proc Natl Acad Sci USA 1992; 89:2027-2031.

16. McClain DA, Maegawa H, Scott Thies R, Olefsky JM. Dissection of the growth versus metabolic effects of insulin and insulin-like growth factor-I in transfected cells expressing kinase-defective human insulin receptors. J Biol Chem 1990; 265:1678-1682.

17. Taylor SI, Cama A, Accili D. Molecular genetics of insulin resistant diabetes mellitus. J Clin Endocrinol Metab 1991; 73:1158-1163.

18. National Diabetes Data Group . Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. Diabetes 1979; 28:1039-1057.

19. DeFronzo RA, Tobin JD, Andres R. Glucose clamp technique: a method for quantifying insulin secretion and resistance. Am J Physiol 1979; 237:E214-E223.

20. Steele R. Influence of glucose loading and of injected insulin on hepatic glucose production. Ann NY Acad Sci 1959; 82:420-430.

21. Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. Detections of polymorphisms of human DNA by gel electrophoresis as single-stranded conformation polymorphisms. Proc Natl Acad Sci USA 1989; 86:2766-2770.

22. Vestergaard H, Bj.o slashed.rb.ae butted.k C, Andersen PH, Bak JF, Pedersen 0. Impaired expression of glycogen synthase mRNA in skeletal muscle of NIDDM patients. Diabetes 1991; 40:1740-1745.

23. Chomczynski P, Sachhi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987; 162:156-159.

24. Hother-Nielsen O, Schmitz O, Bak JF, Beck-Nielsen H. Enhanced hepatic insulin sensitivity, but peripheral insulin resistance in patients with type I (insulin-dependent) diabetes. Diabetologia 1987; 30:834-840.

25. Heding LG. Determination of total serum insulin (IRI) in insulin-treated diabetic patients. Diabetologia 1972; 8:260-266.

26. Heding LG. Radioimmunological determination of human C-peptide in serum. Diabetologia 1975; 11:541-548.

27. White MF, Maron R, Kahn CR. Insulin rapidly stimulates tyrosine phosphorylation of a Mr 185,000 protein in intact cells. Nature (Lond) 1985; 318:183-186.

28. White MF, Stegmann EW, Dull TJ, Ullrich A, Kahn CR. Characterization of an endogenous substrate of the insulin receptor in cultured cells. J Biol Chem 1987; 262:9769-9777.

TABLE 1

Phenotypical characterization of 19 NIDDM patients who were examined with the primary SSCP
scanning of the IRS-1 gene: comparisons to 19 matched controls without detected IRS-1
mutations and 3 glucose-tolerant controls in whom glycine.sup.972 was substituted with arginine

N	Age	Body Mass Index	Plasma HbA.sub.1C	Plasma glucose	Plasma insulin	Plasma C-peptide	M-value	M-value
(F:M)	(yr)	(kg/m.sup.2)	(%)	(mM)	(pM)	(nM)	(basal)	(insulin)
NIDDM								
Mean								
6/13	54	29	7.9	10.4	95	0.92	88	330
.+-.SE								
2 1	0.5	0.9	13	0.09	3	30		
CONTROL (noncarriers of mutation)								
Mean								
7/12	50	28	5.4*	5.5*	60+	0.67.sctn.	77+	484*
.+-.SE								
2 1	0.1	0.1	6	0.04	2	30		
CONTROL (carriers of mutation)								
No. 1								
F	50	27	4.7	5.1	32	0.40	71	461
No. 2								
M	64	23	5.6	6.0	40	0.51	78	389
No. 3								
M	66	23	6.0	5.9	34	0.42	ND	ND

Plasma levels of glucose, insulin and Cpeptide were measured in the fasting state in the morning. Mvalue is the glucose disposal rate (mg/m.sup.2 /min) in the fasting state (basal) in the morning and after 4 hours of euglycemic and hyperinsulinemic clamp (insulin) with steady stat plasma insulin concentrations during the last 30 min of the clamp of 1165 .+-. 71 pM in NIDDM patients and 1034 .+-. 56 pM in controls who were noncarriers of the glycine.sup.972 mutation. In the 3 glucosetolerant mutation carriers the steady state plasma insulin levels were 786 pM (subject no. 1) and 1008 pM (subject no. 2), respectively.
ND = not determined.
*P < 10.sup.-4 vs. NIDDM;
+P < 0.02 vs. NIDDM;
.sctn.P < 0.03 vs. NIDDM.

TABLE 2

Characteristics of 10 NIDDM patients who are carriers of the glycine.sup.972 -> arginine mutation in the IRS-1 gene: comparisons to NIDDM patients who are noncarriers of the mutation

	+Mutation	-Mutation
N (F/M)	3/7	30/46
Age (yr)	52 .+-. 2	54 .+-. 1
Body Mass Index (kg/m.sup.2)	29 .+-. 2	30 .+-. 1
Known duration of diabetes (yr)	5 .+-. 1	4 .+-. 1
HbA.sub.1C (%)	8.6 .+-. 0.6	8.0 .+-. 0.2
Plasma glucose (mM)	12.4 .+-. 1.1	11.7 .+-. 0.5
Plasma insulin (pM)	53 .+-. 10	94 .+-. 8*
Plasma C-peptide (nM)	0.49 .+-. 0.06	0.78 .+-. 0.05 +
M-value (basal)	86 .+-. 5	91 .+-. 3
M-value (insulin)	286 .+-. 29	265 .+-. 16

Value are means .+-. SE. Plasma concentrations of glucose, insulin and Cpeptide were measured in the fasting state in the morning. Mvalue is the glucose disposal rate (mg/m.sup.2 /min) in the fasting state (basal) in the morning and after 4 hours of euglycemic and hyperinsulinemic clamp (insulin) with steady state plasma insulin concentrations during the last 30 min of the clamp of 980 .+-. 57 pM in NIDDM patients who were carriers of the glycine.sup.972 mutation and 1153 .+-. 34 pM in NIDDM patients wh were noncarriers, respectively.
*P < 0.05; +P < 0.02.

TABLE 3

Partial nucloetide (3392-3562) and amino acid (938-994) sequence of the published human
insulin receptor substrate 1 (IRS-1). The glycine to arginine mutation is located at codon
972 (shown in bold). Two YMXM motifs which are putative recognition sites for insulin
signal transmission proteins carrying src homology-2 domains are underlined.

##STR1##
aa938GlyThrGluGluTyrMetLysMetAspLeuGlyProGlyArgArgAlaAlaTrpGln
##STR2##
##STR3##
ATTTGCAGGCCTACCCGGGCAGTGCCCAGCAGCCGGGGTGACTACATGACCATGCAG-3'
3562
TAAACGTCCGGATGGGCCCGTCACGGGTCGTCGGCCCCACTGATGTACTGGTACGTC-5'
IleCysArgProThrArgAlaValProSerSerArgGlyAspTyrMetThrMetGln (SEQ ID NOS: 1 and 2) 994

SEQUENCE LISTING
(1) GENERAL INFORMATION:
(iii) NUMBER OF SEQUENCES: 41
(2) INFORMATION FOR SEQ ID NO: 1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6152 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)
(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 581..4309
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
TGGTATTTGGGCGGCTGGTGGCGGCGGGGACTGTTGGAGGGTGGGAGGAGGCGAAGGAGG60
AGGGAGAACCCCGTGCAACGTTGGGACTTGGCAACCCGCCTCCCCCTGCCCAAGGATATT120
TAATTTGCCTCGGGAATCGCTGCTTCCAGAGGGGAACTCAGGAGGGAAGGGGGCGCGCGC180
TCCTGGAGGGGCACCGCAGGGACCCCCGACTGTCGCCTCCCTGTGCCGGACTCCAGCCGG240
GGCGACGAGAGATGCATCTTCGCTCCTTCCTGGTGGCGGCGGCGGCTGAGAGGAGACTTG300
GCTCTCGGAGGATCGGGGCTGCCCTCACCCCGGACGCACTGCCTCCCCGCCGGGCGTGAA360
GCGCCCGAAAACTCCGGTCGGGCTCTCTCCTGGGCTCAGCAGCTGCGTCCTCCTTCAGCT420
GCCCCTCCCCGCGCGGGGGGCGGCGTGGATTTCAGAGTCGGGGTTTCTGCTGCCTCCAGC480
CCTGTTTGCATGTGCCGGGCCGCGGCGAGGAGCCTCCGCCCCCCACCCGGTTGTTTTTCG540
GAGCCTCCCTCTGCTCAGCGTTGGTGGTGGCGGTGGCAGCATGGCGAGCCCTCCG595
MetAlaSerProPro
15
GAGAGCGATGGCTTCTCGGACGTGCGCAAGGTGGGCTACCTGCGCAAA643
GluSerAspGlyPheSerAspValArgLysValGlyTyrLeuArgLys
101520
CCCAAGAGCATGCACAAACGCTTCTTCGTACTGCGCGCGGCCAGCGAG691
ProLysSerMetHisLysArgPhePheValLeuArgAlaAlaSerGlu
253035
GCTGGGGGCCCGGCGCGCCTCGAGTACTACGAGAACGAGAAGAAGTGG739
AlaGlyGlyProAlaArgLeuGluTyrTyrGluAsnGluLysLysTrp
404550
CGGCACAAGTCGAGCGCCCCCAAACGCTCGATCCCCCTTGAGAGCTGC787
ArgHisLysSerSerAlaProLysArgSerIleProLeuGluSerCys
556065
TTCAACATCAACAAGCGGGCTGACTCCAAGAACAAGCACCTGGTGGCT835
PheAsnIleAsnLysArgAlaAspSerLysAsnLysHisLeuValAla
70758085
CTCTACACCCGGGACGAGCACTTTGCCATCGCGGCGGACAGCGAGGCC883
LeuTyrThrArgAspGluHisPheAlaIleAlaAlaAspSerGluAla
9095100
GAGCAAGACAGCTGGTACCAGGCTCTCCTACAGCTGCACAACCGTGCT931
GluGlnAspSerTrpTyrGlnAlaLeuLeuGlnLeuHisAsnArgAla
105110115
AAGGGCCACCACGACGGAGCTGCGGCCCTCGGGGCGGGAGGTGGTGGT979
LysGlyHisHisAspGlyAlaAlaAlaLeuGlyAlaGlyGlyGlyGly
120125130
GGGGGCAGCTGCAGCGGCAGCTCCGGCCTTGGTGAGGCTGGGGAGGAC1027
GlyGlySerCysSerGlySerSerGlyLeuGlyGluAlaGlyGluAsp
135140145
TTGAGCTACGGTGACGTGCCCCCAGGACCCGCATTCAAAGAGGTCTGG1075
LeuSerTyrGlyAspValProProGlyProAlaPheLysGluValTrp
150155160165
CAAGTGATCCTGAAGCCCAAGGGCCTGGGTCAGACAAAGAACCTGATT1123
GlnValIleLeuLysProLysGlyLeuGlyGlnThrLysAsnLeuIle
170175180
GGTATCTACCGCCTTTGCCTGACCAGCAAGACCATCAGCTTCGTGAAG1171
GlyIleTyrArgLeuCysLeuThrSerLysThrIleSerPheValLys
185190195
CTGAACTCGGAGGCAGCGGCCGTGGTGCTGCAGCTGATGAACATCAGG1219
LeuAsnSerGluAlaAlaAlaValValLeuGlnLeuMetAsnIleArg
200205210
CGGTGTGGCCACTCGGAAAACTTCTTCTTCATCGAGGTGGGCCGTTCT1267
ArgCysGlyHisSerGluAsnPhePhePheIleGluValGlyArgSer
215220225
GCCGTGACGGGGCCCGGGGAGTTCTGGATGCAGGTGGATGACTCTGTG1315
AlaValThrGlyProGlyGluPheTrpMetGlnValAspAspSerVal

230235240245
GTGGCCCAGAACATGCACGAGACCATCCTGGAGGCCATGCGGGCCATG1363
ValAlaGlnAsnMetHisGluThrIleLeuGluAlaMetArgAlaMet
250255260
AGCGATGAGTTCCGCCCTCGCAGCAAGAGCCAGTCCTCGTCCAACTGC1411
SerAspGluPheArgProArgSerLysSerGlnSerSerSerAsnCys
265270275
TCTAACCCCATCAGCGTCCCCCTGCGCCGGCACCATCTCAACAATCCC1459
SerAsnProIleSerValProLeuArgArgHisHisLeuAsnAsnPro
280285290
CCGCCCAGCCAGGTGGGGCTGACCCGCCGATCACGCACTGAGAGCATC1507
ProProSerGlnValGlyLeuThrArgArgSerArgThrGluSerIle
295300305
ACCGCCACCTCCCCGGCCAGCATGGTGGGCGGGAAGCCAGGCTCCTTC1555
ThrAlaThrSerProAlaSerMetValGlyGlyLysProGlySerPhe
310315320325
CGTGTCCGCGCCTCCAGTGACGGCGAAGGCACCATGTCCCGCCCAGCC1603
ArgValArgAlaSerSerAspGlyGluGlyThrMetSerArgProAla
330335340
TCGGTGGACGGCAGCCCTGTGAGTCCCAGCACCAACAGAACCCACGCC1651
SerValAspGlySerProValSerProSerThrAsnArgThrHisAla
345350355
CACCGGCATCGGGGCAGGGCCCGGCTGCACCCCCCGCTCAACCACAGC1699
HisArgHisArgGlyArgAlaArgLeuHisProProLeuAsnHisSer
360365370
CGCTCCATCCCCATGCCGGCTTCCCGCTGCTCCCGTTCGGCCACCAGC1747
ArgSerIleProMetProAlaSerArgCysSerArgSerAlaThrSer
375380385
CCGGTCAGTCTGTCGTCCAGTAGCACCAGTGGCCATGGCTCCACCTCG1795
ProValSerLeuSerSerSerSerThrSerGlyHisGlySerThrSer
390395400405
GATTGTCTCTTCCCACGGCGATCTAGTGCTTCGGTGTCTGGTTCCCCC1843
AspCysLeuPheProArgArgSerSerAlaSerValSerGlySerPro
410415420
AGCGATGGCGGTTTCATCTCCTCGGATGAGTATGGCTCCAGTCCCTGC1891
SerAspGlyGlyPheIleSerSerAspGluTyrGlySerSerProCys
425430435
GATTTCCGGAGTTCCTTCCGCAGTGTCACTCCGGATTCCCTGGGCCAC1939
AspPheArgSerSerPheArgSerValThrProAspSerLeuGlyHis
440445450
ACCCCACCAGCCCGCGGTGAGGAGGAGCTAAGCAACTATATCTGCATG1987
ThrProProAlaArgGlyGluGluGluLeuSerAsnTyrIleCysMet
455460465
GGTGGCAAGGGGCCCTCCACCCTGACCGCCCCCAACGGTCACTACATT2035
GlyGlyLysGlyProSerThrLeuThrAlaProAsnGlyHisTyrIle
470475480485
TTGTCTCGGGGTGGCAATGGCCACCGCTGCACCCCAGGAACAGGCTTG2083
LeuSerArgGlyGlyAsnGlyHisArgCysThrProGlyThrGlyLeu
490495500
GGCACGAGTCCAGCCTTGGCTGGGGATGAAGCAGCCAGTGCTGCAGAT2131
GlyThrSerProAlaLeuAlaGlyAspGluAlaAlaSerAlaAlaAsp
505510515
CTGGATAATCGGTTCCGAAAGAGAACTCACTCGGCAGGCACATCCCCT2179
LeuAspAsnArgPheArgLysArgThrHisSerAlaGlyThrSerPro
520525530
ACCATTACCCACCAGAAGACCCCGTCCCAGTCCTCAGTGGCTTCCATT2227
ThrIleThrHisGlnLysThrProSerGlnSerSerValAlaSerIle
535540545
GAGGAGTACACAGAGATGATGCCTGCCTACCCACCAGGAGGTGGCAGT2275
GluGluTyrThrGluMetMetProAlaTyrProProGlyGlyGlySer
550555560565

GGAGGCCGACTGCCGGGACACAGGCACTCCGCCTTCGTGCCCACCCGC2323
GlyGlyArgLeuProGlyHisArgHisSerAlaPheValProThrArg
570575580
TCCTACCCAGAGGAGGGTCTGGAAATGCACCCCTTGGAGCGTCGGGGG2371
SerTyrProGluGluGlyLeuGluMetHisProLeuGluArgArgGly
585590595
GGGCACCACCGCCCAGACAGCTCCACCCTCCACACGGATGATGGCTAC2419
GlyHisHisArgProAspSerSerThrLeuHisThrAspAspGlyTyr
600605610
ATGCCCATGTCCCCAGGGGTGGCCCCAGTGCCCAGTGGCCGAAAGGGC2467
MetProMetSerProGlyValAlaProValProSerGlyArgLysGly
615620625
AGTGGAGACTATATGCCCATGAGCCCCAAGAGCGTATCTGCCCCACAG2515
SerGlyAspTyrMetProMetSerProLysSerValSerAlaProGln
630635640645
CAGATCATCAATCCCATCAGACGCCATCCCCAGAGAGTGGACCCCAAT2563
GlnIleIleAsnProIleArgArgHisProGlnArgValAspProAsn
650655660
GGCTACATGATGATGTCCCCCAGCGGTGGCTGCTCTCCTGACATTGGA2611
GlyTyrMetMetMetSerProSerGlyGlyCysSerProAspIleGly
665670675
GGTGGCCCCAGCAGCAGCAGCAGCAGCAGCAACGCCGTCCCTTCCGGG2659
GlyGlyProSerSerSerSerSerSerSerAsnAlaValProSerGly
680685690
ACCAGCTATGGAAAGCTGTGGACAAACGGGGTAGGGGGCCACCACTCT2707
ThrSerTyrGlyLysLeuTrpThrAsnGlyValGlyGlyHisHisSer
695700705
CATGTCTTGCCTCACCCCAAACCCCCAGTGGAGAGCAGCGGTGGTAAG2755
HisValLeuProHisProLysProProValGluSerSerGlyGlyLys
710715720725
CTCTTACCTTGCACAGGTGACTACATGAACATGTCACCAGTGGGGGAC2803
LeuLeuProCysThrGlyAspTyrMetAsnMetSerProValGlyAsp
730735740
TCCAACACCAGCAGCCCCTCCGACTGCTACTACGGCCCTGAGGACCCC2851
SerAsnThrSerSerProSerAspCysTyrTyrGlyProGluAspPro
745750755
CAGCACAAGCCAGTCCTCTCCTACTACTCATTGCCAAGATCCTTTAAG2899
GlnHisLysProValLeuSerTyrTyrSerLeuProArgSerPheLys
760765770
CACACCCAGCGCCCCCGGGGAGCCGGAGGAGGGTGCCCGGCATCAGCAC2947
HisThrGlnArgProGlyGluProGluGluGlyAlaArgHisGlnHis
775780785
CTCCGCCTTTCCACTAGCTCTGGTGGCCTTCTCTATGCTGCAACAGCA2995
LeuArgLeuSerThrSerSerGlyGlyLeuLeuTyrAlaAlaThrAla
790795800805
GATGATTCTTCCTCTTCCACCAGCAGCGACAGCCTGGGTGGGGGATAC3043
AspAspSerSerSerSerThrSerSerAspSerLeuGlyGlyGlyTyr
810815820
TGCGGGGCTAGGCTGGAGCCCAGCCTTCCACATCCCCACCATCAGGTT3091
CysGlyAlaArgLeuGluProSerLeuProHisProHisHisGlnVal
825830835
CTGCAGCCCCATCTGCCTCGAAAGGTGGACACAGCTGCTCAGACCAAT3139
LeuGlnProHisLeuProArgLysValAspThrAlaAlaGlnThrAsn
840845850
AGCCGCCTGGCCCGGCCCACGAGGCTGTCCCTGGGGGATCCCAAGGCC3187
SerArgLeuAlaArgProThrArgLeuSerLeuGlyAspProLysAla
855860865
AGCACCTTACCTCGGGCCCGAGAGCAGCAGCAGCAGCAGCCCTTG3235
SerThrLeuProArgAlaArgGluGlnGlnGlnGlnGlnGlnProLeu
870875880885
CTGCACCCTCCAGAGCCCAAGAGCCCGGGGGAATATGTCAATATTGAA3283

LeuHisProProGluProLysSerProGlyGluTyrValAsnIleGlu
890895900
TTTGGGAGTGATCAGTCTGGCTACTTGTCTGGCCCGGTGGCTTTCCAC3331
PheGlySerAspGlnSerGlyTyrLeuSerGlyProValAlaPheHis
905910915
AGCTCACCTTCTGTCAGGTGTCCATCCCAGCTCCAGCCAGCTCCCAGA3379
SerSerProSerValArgCysProSerGlnLeuGlnProAlaProArg
920925930
GAGGAAGAGACTGGCACTGAGGAGTACATGAAGATGGACCTGGGGCCG3427
GluGluGluThrGlyThrGluGluTyrMetLysMetAspLeuGlyPro
935940945
GGCCGGAGGGCAGCCTGGCAGGAGAGCACTGGGGTCGAGATGGGCAGA3475
GlyArgArgAlaAlaTrpGlnGluSerThrGlyValGluMetGlyArg
950955960965
CTGGGCCCTGCACCTCCCAGGGCTGCTAGCATTTGCAGGCCTACCCGG3523
LeuGlyProAlaProProArgAlaAlaSerIleCysArgProThrArg
970975980
GCAGTGCCCAGCAGCCGGGGTGACTACATGACCATGCAGATGAGTTGT3571
AlaValProSerSerArgGlyAspTyrMetThrMetGlnMetSerCys
985990995
CCCCGTCAGAGCTACGTGGACACCTCGCCAGCTGCCCCTGTAAGCTAT3619
ProArgGlnSerTyrValAspThrSerProAlaAlaProValSerTyr
100010051010
GCTGACATGCGAACAGGCATTGCTGCAGAGGAGGTGAGCCTGCCCAGG3667
AlaAspMetArgThrGlyIleAlaAlaGluGluValSerLeuProArg
101510201025
GCCACCATGGCTGCTGCCTCCTCATCCTCAGCAGCCTCTGCTTCCCCG3715
AlaThrMetAlaAlaAlaSerSerSerSerAlaAlaSerAlaSerPro
1030103510401045
ACTGGGCCTCAAGGGGCAGCAGAGCTGGCTGCCCACTCGTCCCTGCTG3763
ThrGlyProGlnGlyAlaAlaGluLeuAlaAlaHisSerSerLeuLeu
105010551060
GGGGGCCCACAAGGACCTGGGGGCATGAGCGCCTTCACCCGGGTGAAC3811
GlyGlyProGlnGlyProGlyGlyMetSerAlaPheThrArgValAsn
106510701075
CTCAGTCCTAACCGCAACCAGAGTGCCAAAGTGATCCGTGCAGACCCA3859
LeuSerProAsnArgAsnGlnSerAlaLysValIleArgAlaAspPro
108010851090
CAAGGGTGCCGGCGGAGGCATAGCTCCGAGACTTTCTCCTCAACACCC3907
GlnGlyCysArgArgArgHisSerSerGluThrPheSerSerThrPro
109511001105
AGTGCCACCCGGGTGGGCAACACAGTGCCCTTTGGAGCGGGGGCAGCA3955
SerAlaThrArgValGlyAsnThrValProPheGlyAlaGlyAlaAla
1110111511201125
GTAGGGGGCGGTGGCGGTAGCAGCAGCAGCAGCGAGGATGTGAAACGC4003
ValGlyGlyGlyGlyGlySerSerSerSerSerGluAspValLysArg
113011351140
CACAGCTCTGCTTCCTTTGAGAATGTGTGGCTGAGGCCTGGGGAGCTT4051
HisSerSerAlaSerPheGluAsnValTrpLeuArgProGlyGluLeu
114511501155
GGGGGAGCCCCCAAGGAGCCAGCCAAACTGTGTGGGGCTGCTGGGGGT4099
GlyGlyAlaProLysGluProAlaLysLeuCysGlyAlaAlaGlyGly
116011651170
TTGGAGAATGGTCTTAACTACATAGACCTGGATTTGGTCAAGGACTTC4147
LeuGluAsnGlyLeuAsnTyrIleAspLeuAspLeuValLysAspPhe
117511801185
AAACAGTGCCCTCAGGAGTGCACCCCTGAACCGCAGCCTCCCCCACCC4195
LysGlnCysProGlnGluCysThrProGluProGlnProProProPro
1190119512001205
CCACCCCCTCATCAACCCCTGGGCAGCGGTGAGAGCAGCTCCACCCGC4243
ProProProHisGlnProLeuGlySerGlyGluSerSerSerThrArg

121012151220
CGCTCAAGTGAGGATTTAAGCGCCTATGCCAGCATCAGTTTCCAGAAG4291

ArgSerSerGluAspLeuSerAlaTyrAlaSerIleSerPheGlnLys
122512301235
CAGCCAGAGGACCGTCAGTAGCTCAACTGGACATCACAGCAGGTCGTT4339
GlnProGluAspArgGln
1240
TCATGGTGACAAAGTCAGAAGACAAAACTGCTTTTAACCTTGTCCTTGAATTCTGTTCTT4399
CGCCTCTGCCCCTTCCTGTTCTTTCCCACTGCTTCCTCAGGGAGAATGCACTTACATTCT4459
CAGGGCATACAAGATGCTCACCCACACTGACATTGGCAGAGAGTCAAACAAACATGTAGG4519
AGCAGCCACAGGAGGGCTTTTTCGTTTGAGGAATTCCCAAGTGAAGTAGTTACTGCAGTA4579
TTTTTAAACATATATCCTATGCCAGTTCTGCGTTTTGTAGAGTTCCTCCGTAAGAAGCTT4639
GATTTGTTTGTTGAAGTTTTCTTTTCACTATATATTTAGGTCAGCCCCTGGAAGGACAGT4699
TCTACAAAAAATACCCGTTAACACAGGGGCTAAACCCTTCCTTATCTTAAACTATCTTAA4759
TAGTTTCTGGGAGCCCTTAAGGGTGATCTTATCAAGTTGTTCTCTGTACTTTTGTTCTGT4819
GATTTCATAATACTAGGGCAACATAAACAGCAGCGGGAAGCATTGATTTCTATTCATCCT4879
GCCCTAAAAAGATCAGGAGTAAGAGCTTTTTAGAAATATGTATTTAGAGAGAAGTACCTA4939
TCTATTTTGTGATCTCTCAAGAAAGTAATTATGGGTGACGTTCTCCTTTTGTTCATGTAC4999
CAGGATTTGTGAAATATTATTCACACAACCGACCCACCATCCCACGGGCCTGGCCTCTCT5059
TGTACAGGATATGCAGGAAACTCTGTATGTGTCTGGGACCCATTATTAAGAGTTATGGGA5119
GTTCATCCTAGGATGTCTGCCTTATAGTTATCTCTTCTTCTTGCACTGAGACATTACAGA5179
TATCATTTGGGGGCTACTATATATCTTCTGTAAAATTACTTTTATTTGTTGAAGAAGAAT5239
GCATACTAAGTCAGGAACATGCCTTAATTTGTTTTGTTTTGCAATTGAGTAGAAGGGCTA5299
AACTGTATCCCTCCACTTTTAGGGTTATTTGCCTGTGTGCCTTTAAGTTCAAAAGTAGAC5359
ACCACAGTAAATGCTGAAAGTTGGCTTTAGGTCTTCTGTGGCTAATGCCGTATTAAAAAT5419
GAAAAACATTTGTGGTAGAAATTAGCCTGCCCTTCGTTCTGTTGATCCTGTTTTCTGGTG5479
GTCATAATTGTGGGTAGAAGAGAGTACAGTTTGCAAATAATGTGATGAGTTGGCAATGCA5539
GAAGTTTCCAGCATTTGGAAACACTTTACTCTGACAAACTGATTATCTTGTGAATTTTAT5599
CTATGCTCCACAGAATGAGCTTTTAAAAGCACTGATTTTTCTTAATTTGTGTCCATTCAT5659
AAGAAATTAATCTGTGCCCTGGTTTCCTATTGACAGGTATTTATTTATCATGTGTTCATA5719
GTCTTCTTAATTCTGTTTCCAATATTTGATCCATATAATTCTCTATTTTATAAAGCAAGA5779
AAAAGGTATATGAACACTCAAATGAAGATTTTGGGTGATATGTTACAAAAAGCATTTATT5839
TGATCAGTATTTACTTCAACATTTATTTTCATCATTCACTAGAAGAAAGATTTAATTGTG5899
TATATCAACATCAGTAGTACAAATCTTGTTATATCAAATGATGTTTTTGGGAGTTCAGAA5959
TCCCTCAACACTTTAAGCATTTGTATTATAAAGTGCCTCATTGGTAAAATAATGAGAATT6019
TGAAGAAAACCAGCCCAGCAGAACTAAAATTTTGGTTTTAAAGGAGATAAAGAGAATAAG6079
TTTTTCTTACTTGTCATCTTAATTTGTTTAGGTTTCTTTTTATAGAGTAGAATAAATGAT6139
GTTTGCTCTGAAG6152
(2) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1243 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
MetAlaSerProProGluSerAspGlyPheSerAspValArgLysVal
151015
GlyTyrLeuArgLysProLysSerMetHisLysArgPhePheValLeu
202530
ArgAlaAlaSerGluAlaGlyGlyProAlaArgLeuGluTyrTyrGlu
354045
AsnGluLysLysTrpArgHisLysSerSerAlaProLysArgSerIle
505560
ProLeuGluSerCysPheAsnIleAsnLysArgAlaAspSerLysAsn
65707580
LysHisLeuValAlaLeuTyrThrArgAspGluHisPheAlaIleAla
859095
AlaAspSerGluAlaGluGlnAspSerTrpTyrGlnAlaLeuLeuGln
100105110
LeuHisAsnArgAlaLysGlyHisHisAspGlyAlaAlaAlaLeuGly

115120125
AlaGlyGlyGlyGlyGlySerCysSerGlySerSerGlyLeuGly
130135140
GluAlaGlyGluAspLeuSerTyrGlyAspValProProGlyProAla
145150155160
PheLysGluValTrpGlnValIleLeuLysProLysGlyLeuGlyGln
165170175
ThrLysAsnLeuIleGlyIleTyrArgLeuCysLeuThrSerLysThr
180185190
IleSerPheValLysLeuAsnSerGluAlaAlaAlaValValLeuGln
195200205
LeuMetAsnIleArgArgCysGlyHisSerGluAsnPhePhePheIle
210215220
GluValGlyArgSerAlaValThrGlyProGlyGluPheTrpMetGln
225230235240
ValAspAspSerValValAlaGlnAsnMetHisGluThrIleLeuGlu
245250255
AlaMetArgAlaMetSerAspGluPheArgProArgSerLysSerGln
260265270
SerSerSerAsnCysSerAsnProIleSerValProLeuArgArgHis
275280285
HisLeuAsnAsnProProProSerGlnValGlyLeuThrArgArgSer
290295300
ArgThrGluSerIleThrAlaThrSerProAlaSerMetValGlyGly
305310315320
LysProGlySerPheArgValArgAlaSerSerAspGlyGluGlyThr
325330335
MetSerArgProAlaSerValAspGlySerProValSerProSerThr
340345350
AsnArgThrHisAlaHisArgHisArgGlyArgAlaArgLeuHisPro
355360365
ProLeuAsnHisSerArgSerIleProMetProAlaSerArgCysSer
370375380
ArgSerAlaThrSerProValSerLeuSerSerSerSerThrSerGly
385390395400
HisGlySerThrSerAspCysLeuPheProArgArgSerSerAlaSer
405410415
ValSerGlySerProSerAspGlyGlyPheIleSerSerAspGluTyr
420425430
GlySerSerProCysAspPheArgSerSerPheArgSerValThrPro
435440445
AspSerLeuGlyHisThrProProAlaArgGlyGluGluGluLeuSer
450455460
AsnTyrIleCysMetGlyGlyLysGlyProSerThrLeuThrAlaPro
465470475480
AsnGlyHisTyrIleLeuSerArgGlyGlyAsnGlyHisArgCysThr
485490495
ProGlyThrGlyLeuGlyThrSerProAlaLeuAlaGlyAspGluAla
500505510
AlaSerAlaAlaAspLeuAspAsnArgPheArgLysArgThrHisSer
515520525
AlaGlyThrSerProThrIleThrHisGlnLysThrProSerGlnSer
530535540
SerValAlaSerIleGluGluTyrThrGluMetMetProAlaTyrPro
545550555560
ProGlyGlyGlySerGlyGlyArgLeuProGlyHisArgHisSerAla
565570575
PheValProThrArgSerTyrProGluGluGlyLeuGluMetHisPro
580585590
LeuGluArgArgGlyGlyHisHisArgProAspSerSerThrLeuHis
595600605

ThrAspAspGlyTyrMetProMetSerProGlyValAlaProValPro
610615620
SerGlyArgLysGlySerGlyAspTyrMetProMetSerProLysSer
625630635640
ValSerAlaProGlnGlnIleIleAsnProIleArgArgHisProGln
645650655
ArgValAspProAsnGlyTyrMetMetMetSerProSerGlyGlyCys
660665670
SerProAspIleGlyGlyGlyProSerSerSerSerSerSerSerAsn
675680685
AlaValProSerGlyThrSerTyrGlyLysLeuTrpThrAsnGlyVal
690695700
GlyGlyHisHisSerHisValLeuProHisProLysProProValGlu
705710715720
SerSerGlyGlyLysLeuLeuProCysThrGlyAspTyrMetAsnMet
725730735
SerProValGlyAspSerAsnThrSerSerProSerAspCysTyrTyr
740745750
GlyProGluAspProGlnHisLysProValLeuSerTyrTyrSerLeu
755760765
ProArgSerPheLysHisThrGlnArgProGlyGluProGluGluGly
770775780
AlaArgHisGlnHisLeuArgLeuSerThrSerSerGlyGlyLeuLeu
785790795800
TyrAlaAlaThrAlaAspAspSerSerSerSerThrSerSerAspSer
805810815
LeuGlyGlyGlyTyrCysGlyAlaArgLeuGluProSerLeuProHis
820825830
ProHisHisGlnValLeuGlnProHisLeuProArgLysValAspThr
835840845
AlaAlaGlnThrAsnSerArgLeuAlaArgProThrArgLeuSerLeu
850855860
GlyAspProLysAlaSerThrLeuProArgAlaArgGluGlnGlnGln
865870875880
GlnGlnGlnProLeuLeuHisProProGluProLysSerProGlyGlu
885890895
TyrValAsnIleGluPheGlySerAspGlnSerGlyTyrLeuSerGly
900905910
ProValAlaPheHisSerSerProSerValArgCysProSerGlnLeu
915920925
GlnProAlaProArgGluGluGluThrGlyThrGluGluTyrMetLys
930935940
MetAspLeuGlyProGlyArgArgAlaAlaTrpGlnGluSerThrGly
945950955960
ValGluMetGlyArgLeuGlyProAlaProProArgAlaAlaSerIle
965970975
CysArgProThrArgAlaValProSerSerArgGlyAspTyrMetThr
980985990
MetGlnMetSerCysProArgGlnSerTyrValAspThrSerProAla
99510001005
AlaProValSerTyrAlaAspMetArgThrGlyIleAlaAlaGluGlu
101010151020
ValSerLeuProArgAlaThrMetAlaAlaAlaSerSerSerSerAla
1025103010351040
AlaSerAlaSerProThrGlyProGlnGlyAlaAlaGluLeuAlaAla
104510501055
HisSerSerLeuLeuGlyGlyProGlnGlyProGlyGlyMetSerAla
106010651070
PheThrArgValAsnLeuSerProAsnArgAsnGlnSerAlaLysVal
107510801085
IleArgAlaAspProGlnGlyCysArgArgArgHisSerSerGluThr

109010951100
PheSerSerThrProSerAlaThrArgValGlyAsnThrValProPhe
1105111011151120
GlyAlaGlyAlaAlaValGlyGlyGlyGlyGlySerSerSerSerSer
112511301135
GluAspValLysArgHisSerSerAlaSerPheGluAsnValTrpLeu
114011451150
ArgProGlyGluLeuGlyGlyAlaProLysGluProAlaLysLeuCys
115511601165
GlyAlaAlaGlyGlyLeuGluAsnGlyLeuAsnTyrIleAspLeuAsp
117011751180
LeuValLysAspPheLysGlnCysProGlnGluCysThrProGluPro
1185119011951200
GlnProProProProProProProHisGlnProLeuGlySerGlyGlu
120512101215
SerSerSerThrArgArgSerSerGluAspLeuSerAlaTyrAlaSer
122012251230
IleSerPheGlnLysGlnProGluAspArgGln
12351240

(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
CCNGG5

(2) INFORMATION FOR SEQ ID NO:4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
GCTCAGCGTTGGTGGTGGCGGTGG24

(2) INFORMATION FOR SEQ ID NO:5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
GCCCGCTTGTTGATGTTGAAGCAGC25

(2) INFORMATION FOR SEQ ID NO:6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
GAAGTGGCGGCACAAGTCGAGCGC24

(2) INFORMATION FOR SEQ ID NO:7:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
CCAGCCTCACCAAGGCCGGAGC22
(2) INFORMATION FOR SEQ ID NO:8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
ACCGTGCTAAGGGCCACCACGACG24
(2) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
AGCACCACGGCCGCTGCCTCC21
(2) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
TCTACCGCCTTTGCCTGACCAGC23
(2) INFORMATION FOR SEQ ID NO:11:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
GCAGTTGGACGAGGACTGGC20
(2) INFORMATION FOR SEQ ID NO:12:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
ACCATCCTGGAGGCCATGCGG21
(2) INFORMATION FOR SEQ ID NO:13:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
AGGGCTGCCGTCCACCGAGGCTGG24
(2) INFORMATION FOR SEQ ID NO:14:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
CAGGCTCCTTCCGTGTCCGCG21
(2) INFORMATION FOR SEQ ID NO:15:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
GAAGCACTAGATCGCCGTGGG21
(2) INFORMATION FOR SEQ ID NO:16:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
CTGTCGTCCAGTAGCACCAGTGG23
(2) INFORMATION FOR SEQ ID NO:17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
GACCGTTGGGGGCGGTCAGGG21
(2) INFORMATION FOR SEQ ID NO:18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
GCGGTGAGGAGGAGCTAAGC20
(2) INFORMATION FOR SEQ ID NO:19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
GCCACTGAGGACTGGGACGGG21
(2) INFORMATION FOR SEQ ID NO:20:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
AGAGAACTCACTCGGCAGGC20
(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
GGCATGTAGCCATCATCCGTGTGG24
(2) INFORMATION FOR SEQ ID NO:22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
ACCCCTTGGAGCGTCGGGGG20
(2) INFORMATION FOR SEQ ID NO:23:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
GCTGGGGCCACCTCCTAAGTCAGG24
(2) INFORMATION FOR SEQ ID NO:24:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
CAGAGAGTGGACCCCAATGG20
(2) INFORMATION FOR SEQ ID NO:25:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
GGCTGCTGGTGTTGGAGTCC20
(2) INFORMATION FOR SEQ ID NO:26:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
TCTTGCCTCACCCCAAACCC20
(2) INFORMATION FOR SEQ ID NO:27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATAGAGAAGGCGACCAGAGC20
(2) INFORMATION FOR SEQ ID NO:28:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
GAGCCGGAGGAGGGTGCCCG20
(2) INFORMATION FOR SEQ ID NO:29:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
AGGTAAGGTGCTGGCCTTGG20
(2) INFORMATION FOR SEQ ID NO:30:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
CAGACCAATAGCCGCCTGGC20
(2) INFORMATION FOR SEQ ID NO:31:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:
TTCATGTACTCCTCAGTGCC20
(2) INFORMATION FOR SEQ ID NO:32:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:
CTTCTGTCAGGTGTCCATCC20
(2) INFORMATION FOR SEQ ID NO:33:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:
TGGCGAGGTGTCCACGTAGC20
(2) INFORMATION FOR SEQ ID NO:34:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:
GGGCAGTGCCCAGCAGCCGG20
(2) INFORMATION FOR SEQ ID NO:35:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
CCAGCAGGGACGAGTGGGCAGC22
(2) INFORMATION FOR SEQ ID NO:36:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:
CCTCAGCAGCCTCTGCTTCC20
(2) INFORMATION FOR SEQ ID NO:37:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
CCACCGCCCCCTACTGCTGC20
(2) INFORMATION FOR SEQ ID NO:38:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
CCACACCCAGTGCCACCCGG20
(2) INFORMATION FOR SEQ ID NO:39:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
GAGGGCACTGTTTGAAGTCC20
(2) INFORMATION FOR SEQ ID NO:40:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
GAGCCAGCCAAACTGTGTGG20
(2) INFORMATION FOR SEQ ID NO:41:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
AACGACCTGCTGTGATGTCC20